REPORT DOCUMENTATION PAGE

	1b. RESTRICTIVE MARKINGS
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 91 0553
6a. NAME OF PERFORMING ORGANIZATION Department of Psychology Yale University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) New, Haven, CT 06520	7b. ADDRESS (City, State and ZIP Code) Bldg 410 Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 86-0357
8c. ADDRESS (City, State and ZIP Code) Bldg 410 Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2313 TASK NO.: A6 WORK UNIT NO.:
11. TITLE (Include Security Classification) Levels of Processing of Speech and Non-Speech	

12. PERSONAL AUTHOR(S)
Arthur G. Samuel

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
FINAL	FROM 9/86 TO 3/90	1991-May-10	20

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The studies conducted examined both signal-dependent factors, and listener-dependent factors. The examinations of signal factors include experiments on perceptual degradation due to signal interruption at critical rates (approximately 4cps), and studies mapping the early levels of representation of speech. The data support the existence of two qualitatively different early processing stages; the first is relatively peripheral and subject to neural fatigue, while the second is central and subject to criterion shifts. The studies of listener based factors include studies of perceptual restoration of deleted sounds (phonemes or musical notes), and studies of the perceptual effect of attentional allocation. The restoration experiments indicate similar architectures in the perceptual processing of speech and music. The attentional investigations demonstrate rather fine-tuned attentional control under high-predictability conditions. Significant progress has been made in achieving the research objective of clarifying the properties of complex auditory pattern recognition.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Dr Genevieve Haddad	(202) 767-5021	NL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

14 MAY 1991

I. Summary

The research project (AFOSR 86-0357) is a broad examination of the perception of complex auditory signals, particularly speech and music. The studies conducted examine both signal-dependent factors, and listener-dependent factors. The examinations of signal factors include experiments on perceptual degradation due to signal interruption at critical rates (approximately 4 cps), and studies mapping the early levels of representation of speech. The data support the existence of two qualitatively different early processing stages; the first is relatively peripheral and subject to neural fatigue, while the second is central and subject to criterion shifts. The studies of listener-based factors include studies of perceptual restoration of deleted sounds (phonemes or musical notes), and studies of the perceptual effect of attentional allocation. The restoration experiments indicate similar architectures in the perceptual processing of speech and music. The attentional investigations demonstrate rather fine-tuned attentional control under high-predictability conditions. Across several of the research efforts, commonalities in the perception of speech and music have been found. Significant progress has been made in achieving the research objective of clarifying the properties of complex auditory pattern recognition.

II. Research Objectives

The objective of the research project is to delineate principles that underlie the perception of complex auditory patterns. The stimuli used are speech and musical patterns of varying complexity. A wide array of experimental procedures and analyses are used to try to determine properties that are true of the perception of complex auditory patterns across stimulus domains. In addition, we also are interested in discovering any principles that are domain specific (e.g., as "categorical perception" has traditionally been claimed to be a principle of perception specific to the speech domain). The various experimental investigations in the project may be broadly grouped into studies of signal-based factors, and studies of listener-based factors. The former group includes experiments that explore how properties of the input signal determine perception, while the latter group includes studies of how listeners' expectations influence perception/performance. The former group primarily focusses on early representations of the signal, and the latter includes higher-level factors (including, but not limited to, attentional influences). The long-term goal of the research is to understand both signal-based and listener-based factors, and their interaction in the perception of complex auditory patterns.

III. Research completed during the funding period

A. Theoretical background for the research completed

The research conducted over the last three years is part of a long-term, broadly-based research program in our laboratory. The goal of this program is to develop our understanding of how humans perceive complex acoustic stimuli such as speech and music. Such an understanding would ultimately be reflected in a fully-specified theory that describes the architecture of the perceptual process. Work in our laboratory and by others has begun to sketch out such a theory; the completed research has expanded our knowledge in this area in several ways. In this section of the report, we will briefly describe the general issues that we are trying to address, summarize some of our prior findings that bear on these issues, and indicate how the completed research has advanced our understanding in this area.

One way to conceptualize perception is as a two-dimensional array of structures and processes. We can think of the horizontal dimension as a variation in stimulus domain, while the vertical dimension reflects different levels of perceptual analysis. In this conception, there are two basic theoretical questions, and an enormous set of empirical ones.

The first theoretical issue involves the definition of stimulus domains: Along the horizontal dimension, are different perceptual processes and structures involved in the perception of different kinds of stimuli? There are trivial cases for which the answer is clearly "yes." For example, if spoken words are the stimuli in one case, and various odors are the stimuli in another, we would not be surprised to find that different perceptual processes are at work. A more interesting and less obvious case might involve spoken words and familiar melodies. A fundamental question for a theory of perception is whether these two sets of stimuli are handled by one set of perceptual processes, or by two disjoint sets. In the literature on speech perception, most researchers have argued for latter; they have claimed that "speech is special" (e.g., Liberman and Mattingly, 1985). In our laboratory, the working assumption is the reverse: We try to account for perceptual phenomena using general principles as much as possible. In this sense, the long-term goal of our work is to develop a psychophysics of complex sounds. Note that this goal suggests that the basic rules governing perception may not be the same for complex acoustic stimuli as for simple ones (such as pure tones). In our view, speech theorists are correct in noting that speech perception often seems to follow different rules than those that govern the perception of simple tones and noises. However, rather than concluding that speech is special, we believe that complex acoustic stimuli may require their own kind of perceptual processing. Thus, the domain division is not between speech and nonspeech; the horizontal break is between very simple acoustic events and more complex ones.

The second major theoretical issue involves the vertical dimension of perceptual levels of analysis. Consider the perception of a simple sentence, or a simple melody. In each case, a number of different levels of analysis can be postulated, with levels varying in their abstraction from the original signal. Thus, we could logically assume that a basic spectral and temporal analysis of each stimulus is a relatively early level of processing; an analysis into phrases might be a later level in each case. A fundamental goal of an information processing analysis is to determine exactly what the levels of analysis are for a given stimulus. Ideally, these levels should be specified logically, psychologically, and neurologically. In addition, a fully developed theory also must specify the nature of information flow in the system. The simplest system type is one with straight "bottom-up" information flow: The results of the earliest level of analysis are passed up to the next, which in turn passes its results to the next, and so on. A slightly more complex system maintains this strict bottom-up ordering, but allows a lower level to continuously feed its ongoing analyses to a higher level (e.g., McClelland's (1979) cascade model). The most complex systems are fully interactive (e.g., McClelland and Elman, 1986) and partially interactive ones (e.g., Samuel, 1981, 1987, 1990). In interactive models, the analyses at more abstract levels (e.g., of phrases) can influence the analyses at lower levels. Partially interactive models differ from fully interactive ones in that the former only allow certain kinds of top-down influences. For example, in Samuel's (1981, 1987, 1989c) partially interactive model, lexical information (i.e., information prestored in a word's representation) can affect the analysis of sublexical information, but higher level analyses (e.g., sentential) cannot.

Over the past decade and a half, we have made progress in addressing both the issue of domain specificity and the issues of informational representation and flow. A few examples should suffice to illustrate this progress and to provide a framework for the completed research. Consider, for example, the well-known phenomenon of categorical perception, often cited as support for the view that speech is special because speech sounds typically produce categorical perception and nonspeech ones do not. Samuel (1977) gave listeners extenseive training on the ABX discrimination task used in studies of categorical perception. After extensive training, listeners demonstrated noncategorical perception of speech -- their within-category discrimination was well above chance. Two aspects of the post-training functions nicely illustrate some of the themes of our research program. First, the discrimination functions reflected Weber's law, illustrating the applicability of general psychophysical properties to speech. Second, even after extensive training, areas of relatively poor discriminability remained, centered around the prototypical stimulus tokens for each phonetic category. Samuel (1977) suggested that listeners perceive each speech stimulus by mapping it onto a prototypical representation, and that tokens that are similar to the prototype all match so well that they cannot be distinguished. Note that this analysis is a representational claim.

Samuel (1982) tested this claim in a selective adaptation study. If listeners really do have a level of analysis that corresponds to a phonetic prototype, then adaptation with a stimulus that matches the prototype would be expected to be more effective than adaptation with a less prototypical token. Samuel (1982) used a pretest to determine each subject's best category exemplar, and conducted adaptation with that token, and other members of the phonetic category. As predicted, the prototype-matching token produced significantly larger adaptation effects than the other tokens, confirming the validity of this representational construct. The research completed under the current grant includes long-term training studies that mirror the technique used by Samuel (1977), but applied to the issue of attentional allocation during speech perception. Our ongoing research also continues the study of perception by inducing perceptual shifts through the adaptation technique. Indeed, we have conducted a program of research in this area that appears to be providing a detailed understanding of the early levels of perceptual analysis.

The adaptation studies that we have conducted under the grant ultimately derive from Samuel and Newport's (1979) adaptation experiments. That work contained several themes that we have pursued in the last few years. Two aspects are particularly relevant. First, Samuel and Newport focussed on the changes in speech identification that could be induced by the repeated presentation of nonspeech sounds. The very existence of such effects underscores the non-domain-specific nature of speech perception. Second, Samuel and Newport used a filtering manipulation to remove spectral overlap of adapters and test syllables, and found that the adaptation effect remained. This finding implies the involvement of an abstract level of representation in the adaptation effects. This issue of level of analysis recurs throughout the completed research, and should be recognizable as one of the fundamental theoretical issues identified above.

Recall that the other aspect of the representational issue involved the nature of communication among the levels of representation (bottom-up, cascaded, top-down, etc.). An ongoing line of research in our laboratory has focussed on this issue; many of our studies in this area have used the phonemic restoration effect (Warren, 1970) to investigate the question. Restoration occurs when listeners report that an utterance sounds intact despite the fact that the experimenter has deleted a portion of the utterance and replaced it with an extraneous sound (e.g., white noise). Samuel (1981) developed a discrimination paradigm to study restoration that allowed signal detection analyses of the phenomenon. These analyses indicated that lexical information directly affected the phonemic level analyses. In contrast, sentential expectations did not produce perceptual effects, leading Samuel (1981) to argue for a partially interactive architecture in which lexical information can have "top-down" effects, but higher level knowledge cannot.

A very recent study in our laboratory has shown that the same kind of domain generality found in the adaptation work seems to be found in phonemic restoration. DeWitt and Samuel (1990) used the discrimination methodology to examine whether perceptual restoration occurs for musical stimuli. The results for music were parallel to those for speech in some detail. For example, melodic predictability, like sentential predictability, did not increase perceptual restoration of replaced notes. In contrast, expectations generated by major scales and by chords did increase restoration of notes replaced in these musical stimuli. The pattern of results across the set of restoration studies suggests that information from well entrenched representations (e.g., words and chords) is used by the perceptual process to restore parts of the signal obliterated by noise; more constructed representations (e.g., sentences and melodies) do not appear to be directly involved in perception.

The example just cited is characteristic of the set of studies that we have completed with the funding provided by AFOSR 86-0357. In several domains, the general research strategy has been to stress the system, and observe its performance when it is producing an output that is objectively nonveridical. The research has continued the investigation of domain-specificity versus domain-generality. Each investigation included tests that cut across the speech-music distinction. To the extent that the phenomena under study all behaved similarly for speech and music, the basis for developing general properties of the perception of complex stimuli have been strengthened.

The studies we have run share a concern with structural issues as well. This is most explicit in the adaptation research: The experiments were designed to test quite specific representational hypotheses. Similarly, the interrupted-signal work derived from an explicitly representational issue: Is the perceptual degradation due to the disruption of processing of a particular unit, the syllable. Our work has rejected this possibility, and demonstrated that it is traceable to quite general processes (e.g., source assignment), and that the phenomenon holds for entirely nonspeech stimuli (e.g., piano melodies).

As noted above, our general theoretical framework yields the two major theoretical issues of domain-specificity and level of analysis, and many empirical issues. Although the completed research has addressed the basic theoretical issues, we should not lose sight of the fact that there is no substitute for solid empirical effort. We believe that the completed research has made a very significant contribution in this respect. Ultimately, theories of perception must rest on an empirical base. The use of interlocking, theoretically motivated manipulations and analyses has provided us with the basis for fleshing out our notions of structure and process.

B. Annotated abstracts of papers based on the funded research

The research conducted with support from AFOSR 86-0357 explored the perception of complex acoustic patterns. The research effort can be thought of as involving four related foci. Two of these emphasize aspects of the signal itself, and two focus on the influence of the listener. Cutting through all of these themes is an exploration of the commonalities in perception of two complex signal domains: speech and music.

1. Selective Adaptation: Our work on selective adaptation effects is aimed at elucidating relatively early levels of analysis of complex sounds. As described above, this work is based on a research program that goes back to work done in the late 1970's. Our previous efforts in this area have identified two qualitatively different levels of analysis. The research done under the current grant has clarified properties of these levels, and suggested that a third level of analysis must exist. This research has appeared in Samuel (1988) and Samuel (1989); recent work, using reaction time methodologies, is being prepared for publication (Samuel and Kat, in preparation).

Abstract from:
Samuel, A.G. (1988). Central and peripheral representation of whispered and voiced speech. Journal of Experimental Psychology: Human Perception and Performance, 14, 379-388.

Whispered speech is very different acoustically from normally voiced speech, yet listeners appear to have little trouble perceiving whispered speech. Two selective adaptation experiments explored the basis for the common perception of whispered and voiced speech, using two synthetic /ba/-/wa/ continua (one voiced, and one whispered). The first experiment used the endpoints of each series as adaptors, and several nonspeech adaptors as well. Speech adaptors produced reliable labeling shifts of syllables matching in periodicity (i.e., whispered-whispered or voiced-voiced); somewhat smaller effects were found with mismatched periodicity. A periodic nonspeech tone with short rise time (the "pluck") produced adaptation effects like those for /ba/. These shifts occurred for whispered test syllables as well as voiced, indicating a common abstract level of representation for voiced and whispered stimuli. Experiment 2 replicated and extended Experiment 1, using same-ear and cross-ear adaptation conditions. There was perfect cross-ear transfer of the nonspeech adaptation effect, again implicating an abstract level of representation. The results support the existence of two levels of processing for complex acoustic signals. The commonality of whispered and voiced speech arises at the second, abstract level. Both this level, and the earlier, more directly acoustic level, are susceptible to adaptation effects.

Abstract from:
Samuel, A.G. (1989). Insights from a failure of selective adaptation: Syllable-initial and syllable-final consonants are different. Perception & Psychophysics, 45, 485-493.

Selective adaptation with a syllable-initial consonant fails to affect perception of the same consonant in syllable-final position, and vice-versa. One account of this well-replicated result invokes a cancellation explanation: With the place of articulation stimuli used, the pattern of formant transitions switches with syllabic position, allowing putative phonetic level effects to be opposed by putative acoustic level effects. Three experiments tested the cancellation hypothesis by preempting the possibility of acoustic countereffects. In Experiment 1, the test syllables and adaptors were /r/-/l/ CVs and VCs which do not produce cancelling formant patterns across syllabic position. In Experiment 2, /b/-/d/ continua were used in a paired-contrast procedure, believed to be sensitive to phonetic, but not acoustic, identity. In Experiment 3, cross-ear adaptation, also believed to tap phonetic rather than acoustic processes, was used. All three experiments refuted the cancellation hypothesis. Instead, it appears that the perceptual process treats syllable-initial consonants and syllable-final ones as inherently different. These results provide support for the use of demisyllabic representations in speech perception.

2. Perceptual degradation: For at least four decades, researchers have periodically interrupted extended utterances to investigate the degradation of perceptual processing (e.g., Cherry and Taylor, 1954; Huggins 1964), or the recovery of performance through additional manipulations, such as the addition of white noise (e.g., Miller and Licklider, 1950). We have conducted a set of experiments that explore the perceptual breakdown that occurs when a message is alternately presented to a listener's right and left ears (using headphones). Huggins (1964) showed that when the alternation rate is about 4 cps, intelligibility drops dramatically; rates under 2 cps, and over 8 cps, produce little degradation. This study also found that if the playback rate was increased by about 20%, the point of minimum performance appeared to shift correspondingly upwards (a result replicated by Wingfield and Wheale, 1975). Huggins' suggestion that the effect is due to disruption of syllabic perceptual units has been widely cited. The work in our laboratory has replicated the basic effect, provided convincing evidence that nonspeech signals show similar degradation through signal alternation, and raised the possibility that a source assignment process (cf Bregman, 1978) may play a role in the phenomenon. Note that the nonspeech results (together with a direct test of the syllabic hypothesis) make it clear that syllabic codes play no role in the phenomenon. In addition, our work suggests that stressing the perceptual system (e.g., by increasing the speech rate) may

selectively disrupt decoding. The research in this area suggests that there are two fundamental perceptual processes that interact to produce the observed perceptual degradation. First, the perceptual system appears to operate in "sampling cycles", triggered by the occurrence of particularly salient events (e.g., the rapid spectral change that occurs at consonant-vowel transitions). Second, the system appears to be interruptable by the apparent onset of a new sound source in the environment. When the rate of such onsets approximates the rate of occurrence of salient acoustic events, perception breaks down. This will occur because each salient event initiates a perceptual sampling cycle, and each such cycle will be interrupted by the occurrence of a new sound source.

<u>Abstract from</u>:
Samuel, A.G. (1991). Perceptual degradation due to signal alternation: Implications for auditory pattern processing. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>17</u>, 392-403.

When a passage is alternately presented to the right and left ears over headphones, perceptual processing is disrupted under certain conditions: When the signal alternation rate is approximately 3-4 cps, intelligibility is greatly reduced. Experiment 1 demonstrated that, contrary to previous theorizing, the effect is not mediated by the disruption of syllabic units. Experiment 2 explored the generality of the perceptual degradation by testing perception of simple piano melodies. The basic effect holds for these complex auditory patterns. The final experiment tested a source-effect explanation of the phenomenon by using three signal locations (right, middle, and left) rather than two. The degree of disruption depends on the likelihood that sounds are assigned to different sources. Together, the experiments help to account for the strikingly selective breakdown in perceptual processing, and speak to the issues of perceptual units, domain-specificity, and auditory source assignment.

<u>3</u>. <u>Perceptual restoration</u>: The perceptual system is designed to produce a filtered version of reality: Incomplete or ambiguous stimuli will usually be perceived as more complete and less ambiguous than the input. The operation of such restoration processes makes it clear that a full understanding of perception must extend beyond the specification of signal-based factors. Studies of perceptual restoration effects, beginning with Warren's (1970) seminal paper, have begun to to clarify the perceptual architecture. The discrimination methodology introduced by Samuel (1981) has proven very useful in distinguishing between perceptual restoration and post-perceptual biases; these results have played an important role in the debate over modular versus interactive architectures (cf Fodor, 1983). Warren replaced part of an utterance with a cough, and found that listeners could not detect the replacement; they appeared to have restored the missing speech. In Samuel's discrimination methodology, stimulus items are constructed in pairs: a <u>replacement</u> item is comparable to Warren's stimuli -- a portion of the waveform is replaced with an extraneous sound (white noise). An <u>added</u> item is constructed by adding the white noise to the same portion of the waveform that is replaced in the matching item. To the extent that listeners are perceptually restoring the missing sound in replacement items, they should sound like added items (intact with an extraneous noise). By using signal detection analyses, a bias-free measure of how much replacement items sound like intact ones is computed (d'), and is the measure of the perceptual strength of the effect; a bias parameter (Beta) is also computed that reflects postperceptual bias toward calling a stimulus intact.

The research in this area supported by the grant examined three general questions. First, what role do pre-existing lexical representations play in generating the restored percept? In particular, two experiments (Samuel, 1987) examined differences in the strength of the phenomenon as a function of the test word's relationship to all words in the listener's mental lexicon. The question at issue was whether restoration is affected by the extent to which a single unique lexical entry is consistent with the incoming sensory data. The second line of research in this area (DeWitt and Samuel, 1990) examined perceptual restoration of musical sounds. As mentioned above, this work was notable for the extent to which the results in the domain of music were parallel to those in speech. The data suggest that there are common processes that operate in the two stimulus domains. The third set of experiments in this area examined whether listeners can attain control over the restoration process via training or attentional cues. This research is summarized below, in the section reviewing the role of attention in speech perception.

Abstract from:
Samuel, A.G. (1987). Lexical uniqueness effects on phonemic restoration. Journal of Memory and Language, 26, 36-56.

Phonemic restoration is a powerful auditory illusion in which listeners hear a part of a word that has in fact been replaced by another sound. Two experiments explore whether the strength of the illusion is affected by whether a single lexical item could be restored. In Experiment 1, more perceptual restoration was found for stimuli that were multiply restorable (e.g., "_egion" -> "legion" or "region") than for lexically unique ones (e.g., "_esion" -> "lesion"). In Experiment 2, lexical uniqueness was examined as a function of time: Words become lexically unique when enough has been heard to eliminate all alternatives. This manipulation also affected the strength of the illusion. The results complement those of other techniques in supporting an active role for lexical representations in the perception of speech.

Abstract from:
DeWitt, L.A., and Samuel, A.G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. Journal of Experimental Psychology: General, 119, 123-144.

When presented with an incomplete signal, the perceptual process attempts to reconstruct the original pattern. With linguistic stimuli, there is now a body of evidence that demonstrates the perceptual involvement of lexical representations; information about a particular word has an on-line perceptual effect (see, e.g., Samuel, 1981, 1987). In contrast, sentential predictability does not seem to be used in perceptual processing; any effects appear to be attributable to post-perceptual decision processes. The results of Experiments 1-5 paint a similar picture in the domain of music. The first three experiments tested several knowledge sources that could potentially be used by the perceptual process: melodic familiarity, rhythmic predictability, and tonality. In each case, the information was not used to increase the restoration of missing notes. Instead, just as with sentences, listeners were able to listen more analytically when more cues of this sort were available. In Experiments 4 and 5, we turned to musical structures whose representations may be more "entrenched", or unitary, than those of melodies. For both scales and chords, we found evidence of perceptual involvement; in both cases, when a manipulation was devised to increase the activation of such a

representation, more perceptual restoration was observed. The chord results suggest that these activations are mediated by the establishment of a sense of key.

We thus find that musical elements that form essentially invariant wholes appear to be accessed during perceptual processing, just as essentially invariant lexical items are. Musical elements that are more "constructed" -- melodies -- do not demonstrate this property, just as constructed sentences do not. We are led to suggest that this is a general property of the perception of complex acoustic stimuli: "Low-level" representations will play a role in perception, while "higher-level" ones will not.

4. Attentional effects in speech perception: A description of the operation of an information processing system must include a characterization of the forms of representation used, the nature of the information flow, and the mechanisms that control processing. In the domain of speech perception, there is a large literature devoted to the representational issue, with evidence adduced for acoustic features (e.g., Sawusch, 1977), phonemes (e.g., Norris and Cutler, 1988), demisyllables (e.g., Fujimura, 1976), syllables (e.g., Massaro, 1975), and words (e.g., Samuel, 1986). There is also a substantial literature that focusses on the question of information flow. Research in this area is concerned with whether the information develops in a strictly "bottom-up", or "autonomous" way (e.g., Cutler and Norris, 1979; Cutler, Mehler, Norris, and Segui, 1987; Massaro, 1989), or whether higher level knowledge can influence the operation of lower-level analyses (e.g., Elman and McClelland, 1984, 1988; Marslen-Wilson and Welsh, 1978; McClelland and Elman, 1986). There are also partially interactive models in which top-down influences are allowed, but only among a subset of the levels in the system (e.g., Connine and Clifton, 1987; Samuel, 1981, 1986; Tanenhaus and Lucas, 1987).

The literature dealing with control processes in speech perception is considerably sparser than the literatures on representation and information flow (see Nusbaum and Schwab, 1986). However, control issues are playing a growing role in assessing models of speech perception. For example, in Cutler and Norris's (1979) autonomous model, decisions in experimental tasks (such as phoneme monitoring) can be based on information encoded at either a phonemic or a lexical level. Control processes are invoked in the model to account for observed differences in performance across testing conditions: Task parameters such as word length, secondary tasks, and stimulus quality can bias attentional allocation toward either lexical or phonemic representations (Cutler et al, 1987). In fact, proponents of the autonomous model have argued that the model's inclusion of this control structure allows it to account for some empirical findings better than interactive models without such control processes (e.g., the TRACE model of McClelland and Elman, 1986).

The most promising route for providing a basis to flesh out the details of attention's role and mechanisms will probably come from developing converging methodologies. Toward this end, we have used two quite different methodologies to explore the role of attention in speech perception. One is the restoration paradigm discussed above. The other is the phoneme monitoring technique. The phoneme monitoring technique continues to be widely used. This paradigm has provided much of the evidence for autonomous theories (e.g., Cutler et al, 1987; Frauenfelder, Segui, and Dijkstra, 1990). The method has also been adopted to explore attentional allocation. For example, target detection times are faster for phonemes in syllables that are expected to be stressed (Shields, McHugh, and Martin, 1974), even when local acoustic cues for stress have been eliminated

(Cutler, 1976; Pitt and Samuel, 1990a). With the elimination of an acoustic basis to the advantage, the most parsimonious account of the effect is that listeners differentially allocate processing resources to aspects of the signal that are expected to provide the most stable analysis; stressed syllables also tend to coincide with high information content words. Most recently, we (Pitt and Samuel, 1990b) have incorporated cost-benefit analyses (Posner, 1980) into the phoneme monitoring technique in order to develop an "attentional profile" during speech perception. In this version of the task, probability manipulations are used to induce subjects to expect targets at certain points in an utterance. Detection times for targets occuring at either an expected location or at an unexpected location were compared to detection times in a control condition in which no localized expectations were induced. The data revealed substantial benefits in reaction time to targets in expected locations, and substantial costs for targets in unexpected locations. Moreover, the technique provides a phoneme-by-phoneme profile of attentional allocation, a profile that indicated very fine tuning of expectations, particularly for word-initial and word-final phonemes. This procedure seems very promising as a converging operation with the techniques used in the restoration work. These procedures show promise of helping to delineate both the flow of information during speech perception, and the processes that control this flow.

<u>Abstract from</u>:
Pitt, M.A., and Samuel, A.G. (1990). The use of rhythm in attending to speech. <u>Journal of Experimental Psychology</u>: <u>Human Perception and Performance</u>, <u>16</u>, 564-573.

Three experiments examined attentional allocation during speech processing to determine whether listeners capitalize on the rhythmic nature of speech and attend more closely to stressed than to unstressed syllables. Ss performed a phoneme monitoring task in which the target phoneme occurred on a syllable that was either predicted to be stressed or unstressed by the context preceding the target word. Stimuli were digitally edited to eliminate the local acoustic correlates of stress. A sentential context and a context composed of word lists, in which all the words had the same stress pattern, were used. In both cases, the results suggest that attention may be preferentially allocated to stressed syllables during speech processing. However, a normal sentence context may not provide strong predictive cues to lexical stress, limiting the use of the attentional focus.

<u>Abstract from</u>:
Pitt, M.A., and Samuel, A.G. (1990). Attentional allocation during speech perception: How fine is the focus? <u>Journal of Memory and Language</u>, <u>29</u>, 611-632.

A variant of the phoneme monitoring task was developed to investigate temporal selective attention during speech processing. In this version of the task the probable location of the target phoneme to be monitored for was varied to induce subjects to attend more closely to one location than to others. Experiments 1 and 2 examined selective attention under normal listening conditions, and Experiment 3 investigated attention under more difficult monitoring conditions. Overall, the data indicate that temporal selective attention is very flexible and precise: Benefits in performance were obtained at the attended location, and costs were observed at the unattended locations. Imposing extra processing demands on the subjects resulted in a loss of attentional selectivity under some circumstances. The implications of these results for issues concerning prelexical processing are also discussed.

Abstract from:
Samuel, A.G. (1991). A further examination of the role of attention in the phonemic restoration illusion. Quarterly Journal of Experimental Psychology, 43A, xxx-xxx.

Models of how listeners understand speech must specify the types of representations that are computed, the nature of the flow of information, and the control structures that modify performance. Three experiments are reported that focus on the control processes in speech perception. Subjects in the experiments tried to discriminate stimuli in which a phoneme had been replaced with white noise from stimuli in which white noise was merely superimposed on a phoneme. In the first two experiments, subjects practiced the discrimination for thousands of trials, but did not improve, suggesting that they have poor access to low-level representations of the speech signal. In the third experiment, each (auditory) stimulus was preceded by a visual cue that could potentially be used to focus attention in order to enhance performance. Only subjects who received information about both the identity of the impending word and the identity of the critical phoneme showed enhanced discrimination. Other cues, including syllabic plus phonemic information, were ineffective. The results indicate that attentional control of processing is difficult but possible, and that lexical representations play a central role in the allocation of attention.

References

Bregman, A.S. (1978). The formation of auditory streams. In J. Requin (ED.), Attention and Performance VII, 63-75. Hillsdale: Lawrence Eribaum.

Cherry, E. C., & Taylor, W. K. (1954). Some further experiments on the recognition of speech with one and with two ears. Journal of the Acoustical Society of America, 26, 554-559.

Connine, C.M. and Clifton, C., Jr. (1987) Interactive use of lexical information in speech perception. Journal of Experimental Psychology: Human Perception and Performance, 13 291-299

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. Perception & Psychophysics, 20, 55-60.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. Cognitive Psychology, 19, 141-177.

Cutler, A., and Norris, D. (1979). Monitoring sentence comprehension. In W.E. Cooper & E.C.T. Walker (Eds.). Sentence processing: Psycholinguistic studies presented to Merrill Garrett (pp. 113-134). Hillsdale, NJ: Erlbaum.

DeWitt, L., and Samuel, A.G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. Journal of Experimental Psychology: General, 119, 123-144.

Elman, J.L., and McClelland, J.L. (1984). Speech as a cognitive process: The interactive activation model. In N. Lass (Ed.), Speech and Language (Vol 10). New York: Academic Press.

Elman, J.L., McClelland, J.L. (1988) Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. Journal of Memory and Language, 27, 143-165.

Fodor, J.A. (1983). The Modularity of Mind. MIT Press, Cambridge, Mass.

Frauenfelder, U.H., Segui, J., and Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? Journal of Experimental Psychology: Human Perception and Performance, 16, 77-91.

Fujimura, O. (1976). Syllables as concatenated demisyllables and affixes. Journal of the Acoustical Society of America, S55.

Huggins, A.W.F. (1964). Distortion of the temporal pattern of speech: Interruption and alternation. Journal of the Acoustical Society of America, 36, 1055-1064.

Liberman, A.L., and Mattingly, I.G. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.

Massaro, D. W. (1975). Preperceptual images, processing time, and perceptual units in speech perception, in Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics, edited by D.W. Massaro (Academic Press, New York).

Massaro, D.W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. Cognitive Psychology, 21, 398-421.

McClelland, J.L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. Psychological Review, 86, 287-330.

McClelland, J.L., and Elman, J.L. (1986). The TRACE model of speech perception. Cognitive Psychology, 18, 1-86.

Miller, G.A., & Licklider, J.C. (1950). The intelligibility of interrupted speech. Journal of the Acoustical Society of America, 22, 167-173.

Norris, D., and Cutler, A. (1988). The relative accessibility of phonemes and syllables. Perception & Psychophysics, 43, 541-550.

Nusbaum, H.C., and Schwab, E.C. (1986). The role of attention and active processing in speech perception. In E.C. Schwab and H.C. Nusbaum (Eds.), Pattern recognition by humans and machines: Speech perception, Vol I. (pp. 113-157). New York: Academic Press.

Pitt, M.A., and Samuel, A.G. (1990). The use of rhythm in attending to speech. Journal of Experimental Psychology: Human Perception and Performance, 16, 564-573.

Pitt, M.A., and Samuel, A.G. (1990). Attentional allocation during speech perception: How fine is the focus? Journal of Memory and Language, 29, 611-632.

Posner, M.I. (1980) Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3-25.

Samuel, A.G. (1977). The effect of discrimination training on speech perception: Noncategorical perception. Perception & Psychophysics, 22, 321-330.

Samuel, A.G. (1981a). Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology: General, 110, 474-494.

Samuel, A.G. (1982). Phonetic prototypes. Perception & Psychophysics, 31, 307-314.

Samuel, A. G.(1986). The role of the lexicon in speech perception. In E. C. Schwab & H. C. Nusbaum (Eds.), Pattern recognition by humans and machines: Speech perception, Vol. I. (pp. 89-112). New York: Academic Press.

Samuel, A.G. (1987). The effect of lexical uniqueness on phonemic restoration. Journal of Memory and Language, 26, 36-56.

Samuel, A.G. (1988). Ear alternation effects are not due to syllabic disruption. (in preparation).

Samuel, A.G. (1989). Insights from a failure of selective adaptation:

Syllable-initial and syllable-final consonants are different. Perception and Psychophysics, 45, 485-493.

Samuel, A.G. (1990). Using perceptual restoration effects to explore the architecture of perception. Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives, G. Altman (Ed.). Cambridge, Mass: MIT Press/Bradford Books.

Samuel, A.G., and Newport, E.L. (1979). Adaptation of speech by nonspeech: Evidence for complex acoustic cue detectors. Journal of Experimental Psychology: Human Perception and Performance, 5, 563-578.

Sawusch, J. (1977). Peripheral and central processes in selective adaptation of place of articulation in stop consonants. Journal of the Acoustical Society of America, 62, 738-750.

Shields, J.L., McHugh, A., and Martin, J.G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. Journal of Experimental Psychology, 102, 250-255.

Tanenhaus, M.K., and Lucas, M.M. (1987). Context effects in lexical processing. Cognition, 25, 213-234.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds. Science, 167, 392-393.

Wingfield, A., and Wheale, J. (1975). Word rate and intelligibility of alternated speech. Perception & Psychophysics, 18, 317-320.

IV. List of Publications

Samuel, A.G. (1987). Lexical uniqueness effects on phonemic restoration. Journal of Memory and Language, 26, 36-56.

Samuel, A.G. (1988). Central and peripheral representation of whispered and voiced speech. Journal of Experimental Psychology: Human Perception and Performance, 14, 379-388.

Samuel, A.G. (1989). Insights from a failure of selective adaptation: Syllable-initial and syllable-final consonants are different. Perception & Psychophysics, 45, 485-493.

Pitt, M.A., and Samuel, A.G. (1990). The use of rhythm in attending to speech. Journal of Experimental Psychology: Human Perception and Performance, 16, 564-573.

DeWitt, L.A., and Samuel, A.G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. Journal of Experimental Psychology: General, 119, 123-144.

Pitt, M.A., and Samuel, A.G. (1990). Attentional allocation during speech perception: How fine is the focus? Journal of Memory and Language, 29, 611-632.

Samuel, A.G. (Ed.). (1990). Modularity of speech and language. In Modularity and the Motor Theory of Speech Perception, I.G. Mattingly and M. Studdert-Kennedy (Eds.). Hillsdale, N.J.: Erlbaum.

Samuel, A.G. (1990). Using perceptual restoration effects to explore the architecture of perception. In Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives, G. Altmann (Ed.). Cambridge, Mass: MIT Press/ Bradford Books.

Samuel, A.G. (1991). Perceptual degradation due to signal alternation: Implications for auditory pattern processing. Journal of Experimental Psychology: Human Perception and Performance, 17, 392-403.

Samuel, A.G. (1991). A further examination of the role of attention in the phonemic restoration illusion. Quarterly Journal of Experimental Psychology, 43A, xxx-xxx.

V. Personnel

Principal Investigator: Arthur G. Samuel, Associate Professor of Psychology at Yale University. Ph.D. from University of California, San Diego, 1979.

Associate in Research: Donna Kat. B.S. in Psychology from University of California. San Diego, 1979.

Graduate Student: Lucinda DeWitt. Ph.D. awarded by Yale University, May 1989. Title: The Role of Knowledge-based Expectations in Music Perception. Dr. DeWitt is now an Assistant Professor of Psychology at Depauw University.

Graduate Student/Post Doc: Mark Pitt. Ph.D. awarded by Yale University, May 1990. Title: The Allocation of Attention During the Processing of Speech. Dr. Pitt is now an Assistant Professor of Psychology at Ohio State University.

VI. List of Interactions (coupling activities)

Conference Presentations

During this reporting period, we have made fifteen presentations at eleven conferences. These are summarized here:

A. Samuel. (Fall 1986) Do Ear-Switching Results Support the Syllable as a Processing Unit? Presented at the Meeting of the Psychonomic Society. New Orleans.

A. Samuel and L. DeWitt. (Fall 1986) Perceptual restoration of musical notes. Presented to the Acoustical Society of America. Anaheim.

A. Samuel (Fall 1986) Timbre: A better musical analogy to speech? Presented to the Acoustical Society of America. Anaheim.

A. Samuel. (Fall 1987) Central and peripheal representation of whispered and voiced speech. Presented to the Acoustical Society of America. Miami.

A. Samuel. (May 1988) Using perceptual restoration effects to explore the architecture of perception. Presented at the Workshop on Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives. Sperlonga, Italy.

A. Samuel. (June 1988) Modularity of speech and language. Presented at the Conference on Modularity and the Motor Theory of Speech Perception. New Haven, CT.

A. Samuel and L. DeWitt. (Fall 1988) Perceptual restoration of musical notes. Presented at the Meeting of the Psychonomic Society. Chicago.

A Samuel. (Fall 1988) Cross-syllabic-position failures of adaptation are not due to acoustic-phonetic cancellation. Presented to the Acoustical Society of America. Honolulu.

D. Kat and A. Samuel. (Fall 1988) Using reaction times to explore the mechanisms of selective adaptation. Presented to the Acoustical Society of America. Honolulu.

A. Samuel (Fall 1989) Effects of Training and Attentional Cues of Phonemic Restoration. Presented to the Psychonomic Society. Atlanta.

M. Pitt and A. Samuel (Fall 1989) Attentional allocation during phoneme monitoring: An investigation into the unit of perceptual analysis and selective attention during speech perception. Presented to the Acoustical Society of America. St. Louis.

A. Samuel and D. Kat. (Fall 1990) Selective Adaptation Five Years Later:

Bringing Home the Herring. Presented to the Psychonomic Society. New Orleans.

M. Pitt and A. Samuel. (Fall 1990) Is Phoneme Identification Facilitated by Feedback from a Words' Lexical Representation? Presented to the Psychonomic Society. New Orleans.

A Samuel. (Fall 1990) Signal alternation disrupts perception of music as well as speech. Presented to the Acoustical Society of America. San Diego.

D. Kat (Fall 1990) Reaction time and dichotic evidence for central processing of complex auditory signals. Presented to the Acoustical Society of America. San Diego.

Other Interactions:

In addition to the conferences listed above, the PI has been heavily involved in the review process for both granting agencies and journals. This includes grant reviews for AFOSR and NSF, and manuscript reviews for:

- Behavioral Research Methods and Instrumentation
- Cognition
- Cognitive Psychology
- Journal of the Acoustical Society of America
- Journal of Experimental Psychology: Human Perception and Performance
- Language and Speech
- Memory and Cognition
- Perception & Psychphysics
- Quaterly Journal of Experimental Psychology

During the granting period, the PI was asked (and agreed) to join the Editorial Board of three journals: Cognition, Memory and Cognition, and the Journal of Experimental Psychology: Human Perception and Performance. These professional activities produce a great deal of interaction with other researchers.